© Academy of Management Learning & Education, 2012, Vol. 11, No. 2, 187–206. http://dx.doi.org/10.5465/amle.2011.0032

Relationship Conflict Improves Team Performance Assessment Accuracy: Evidence From a Multilevel Study

NICOLA BREUGST
HOLGER PATZELT
Technische Universität München

DEAN A. SHEPHERD
HERMAN AGUINIS
Indiana University

This article demonstrates the applicability of implementing multilevel design, measurement, and analysis to explore research questions central to the management education literature. Specifically, we offer a multilevel model of self-assessed team performance based on self-enhancement theory and on the construal level theory of psychological distance. We implemented a team decision-making task with 52 student teams, objectively measured team performance, and then asked team members and teams for assessments of team performance. Results show that relationship conflict moderates the relationship between (a) team objective performance and individual assessments of team performance (i.e., cross-level interaction effect); and (b) team objective performance and team assessments of team performance (i.e., team-level interaction effect). The moderating effect is isomorphic across levels of analysis: As relationship conflict increases, the relationship between subjective and objective performance also increases. Our results have important implications for research on management education, assessments of team performance, relationship conflict, and methodological approaches adopted by management education researchers.

Teamwork plays an important role in both organizational life and education (Chen, Donahue, & Klimoski, 2004; Mathieu, Maynard, Rapp, & Gilson, 2008). For example, teamwork is an important part of entrepreneurship courses and entrepreneurship in practice (Boni, Weingart, & Evenson, 2009), and MBA programs use cooperative learning in teams to train students and help them develop the interpersonal skills necessary to be effective managers (Baldwin, Bedell, & Johnson, 1997). Despite the potential for diffusion of responsibility in a team, and some members who might "free ride" by not contributing to team performance (Willcoxson, 2006), assessing team performance can serve as corrective feedback to teams to help them adjust to the requirements of a task (Schraw, Potenza, & Nebelsick-Gullet, 1993); create a feeling of interdependence; and enhance the members' accountability for team outcomes (Brown, 2003; Goltz, Hietapelto, Reinsch, & Tyrell, 2008; Tonn & Milledge, 2002; Volkema, 2010). Thus, it is not only important that educators help team members to see each other's contribution as has been recently suggested (cf. Freeman & Greenacre, 2011), but also, to help teams develop an accurate assessment of their performance.

However, previous research has shown that it is difficult for people to accurately assess their own performance (Aguinis, Joo, & Gottfredson, 2011; Schraw et al., 1993). For example, based on a meta-

analysis of 166 studies, Sitzmann, Ely, Brown, and Bauer (2010) concluded that learners (in education and in the workplace) generally have difficulty self-assessing their knowledge, competence, or performance in cognitive tasks, such as learning. Specifically, 56% of studies focusing on the accuracy of self-assessment concluded that people assess their own performance imprecisely. Although research on self-assessment at the individual level has made an important contribution to the literature (and will likely continue to do so), given the prevalence of teamwork in education and work settings (Boni et al., 2009; Chen et al., 2004; Stevens & Campion, 1994), there is a need to investigate self-assessments at the team level.

A possible explanation for this surprising gap in the literature is the methodological challenges associated with team and multilevel research. These challenges include potential confounds associated with team formation—self-selection, survivor, or retrospective reporting bias; the nature of the team—team size, previous successes or failures, collective belief, and identification; and the nature of the task—its difficulty and complexity, the distribution of information among members, and time. Research at the team level also involves additional work in collecting data because sample size (and associated degrees of freedom) is determined not only by the number of individuals but also by the number of teams. In fact, management education scholars have not, for the most part, implemented multilevel designs that are needed to address important questions, answers, and theoretical insights. For example, of the 113 published primary-level studies included in the quantitative review by Sitzmann and colleagues (2010), only 12 included measures at different levels of analysis. Moreover, only one of these 113 studies tested hypotheses regarding cross-level interaction effects (i.e., the extent to which a relationship between two variables at one level of analysis changes as values of a variable at another level of analysis changes).

The purpose of this paper is twofold: to demonstrate the applicability of multilevel design, measurement, and analysis to explore team and cross-level research questions central to the management education literature, and to gain a deeper understanding of the accuracy of self-assessments of team performance by considering the team context. Building on research concerning teams and performance assessment, we develop a model of students' self-assessments of team performance that includes relationship conflict—tension, animosity, and friction between team members (Jehn, 1995)—as an important moderator in the relationship between objective performance and self-assessed performance. To test this model, we implemented a team decision-making task (hidden-profile task) with 52 student teams and asked team members and teams for assessments of team performance.

This study makes three primary contributions. First, most research on self-assessment has been survey based and at the individual level of analysis. As described earlier, hypothesis tests involving cross-level interaction effects are virtually absent in the management education literature. Because variables measured at the individual level of analysis (e.g., student learning, student performance) are affected by variables both at the individual and also at the group level of analysis, not implementing multilevel design, measurement, and analysis techniques precludes management education scholars from understanding how and why cross-level interaction effects impact some of the most critical dependent variables in our field. Interaction effects are crucial for the advancement of all scientific fields because they serve as indicators of a theory's boundary conditions (i.e., conditions under which certain relationships may become weaker or stronger or even change in direction, Aguinis, 2004). Moreover, cross-level interaction effects refer to situations when, for example, relationships between variables at the individual student level of analysis may change as we move from one team to another, from one classroom to another, from one instructor to another, or from one university to another (Mathieu, Aguinis, Culpepper, & Chen, in press). In our work here, we use a multilevel research design that includes a hidden-profile decision-making task (Stasser & Titus, 1985). By using this methodological approach to address a research question of high importance to the management education literature, we are able to highlight the advantages, limitations, and future research possibilities of team studies. We do not claim that our approach is a "silver bullet" that overcomes all methodological challenges or limitations, but we believe that it is an important methodological tool for management education scholars to add to their repertoire. For example, although there are questions of generalizability of this study's findings (as with all studies involving substantial experimenter control), we have been able to provide evidence of cause-and-effect relationships at multiple levels of analysis, precise control of variables, and a standardized procedure and method for replication and extension.

Second, we contribute to the literature on the assessment of team performance. By taking into

account the team interaction process, we explain heterogeneity in the accuracy of self-assessments of team performance. We theorize and find that team members' perception of relationship conflict is connected to a state of mind that facilitates accurate assessments of team performance from both the members' and the team's perspectives.

Finally, we contribute to the literature on team conflict. Most studies have found that relationship conflict diminishes task performance (Jehn, 1995; Langfred, 2007; Mohammed & Angell, 2004). However, here, we highlight an important exception: Relationship conflict actually improves team performance when the task consists of assessing team performance. Our results show that at the individual and team levels, relationship conflict enhances the accuracy of team performance assessments.

THEORY DEVELOPMENT AND HYPOTHESES

In our multilevel model of students' self-assessments of team performance, we focus on questions of how well and when students' and student teams' performance assessments reflect objective team performance. We suggest that the relationship between objective performance and individuals' assessments of team performance is influenced by their perception of any relationship conflict experienced during the task. Further, the relationship between objective performance and teams' assessments is influenced by their collective perception of relationship conflict. Building on the team literature, we develop specific hypotheses about each effect in the subsequent sections.

Students' Assessments of Team Performance

We build on work in organizational contexts and define teams as "a distinguishable set of two or more people who interact, dynamically, interdependently, and adaptively toward a common and valued goal/objective/mission, who have each been assigned specific roles or functions to perform, and who have a limited life-span of membership" (Salas, Dickinson, Converse, & Tannenbaum, 1992: 4). To achieve the team's collective outcome, all members need to contribute to their team's task (Cohen & Bailey, 1997). For example, when the team task is decision making, the team can achieve the optimal answer only when all of the members share their unique information (Schulz-Hardt, Brodbeck, Mojzisch, Kerschreiter, & Frey, 2006; Stasser & Titus, 1985). For a number of tasks, there are objective criteria that indicate how well a team performed; for example, the amount of coal mined by a team (Goodman & Leyden, 1991), the correct assembly of devices (Lewis, Belliveau, Herndon, & Keller, 2007), or the sales performance of teams in retail stores (George & Bettenhausen, 1990). However, members may not always be aware of the objective performance of their team. Indeed, it is often difficult for people to accurately assess their own performance (Aguinis et al., 2011; Schraw et al., 1993). For instance, in a recent meta-analysis, Sitzmann and colleagues (2010) found that the mean-corrected correlation of students' self-assessments of knowledge, competence, or performance in cognitive tasks and their actual performance was $\rho = .34$. In other words, student self-assessments accounted for only 11.56% of variance in actual performance. Stated differently, almost 90% of variance in student ratings of their own performance was unrelated to actual performance.

The divergence of perceived performance assessments from objective performance can be explained by self-enhancement theory and the construal-level theory of psychological distance. First, according to self-enhancement theory (Allport, 1937), individuals want to achieve or maintain a positive image of the self and increase self-esteem. For example, people believe that they are responsible for success but not for failure (self-serving bias, Bradley, 1978), or they think that their current self is better than past selves (Wilson & Ross, 2001). This tendency to self-enhance also occurs in social environments; that is, individuals tend to rate their own group more positively than members from out-groups (Rabbie & Horwitz, 1969). This effect applies to teams (Bettencourt & Dorr, 1998) and even occurs in a minimal group situation in which membership is based on an irrelevant criterion, the group only exists for a short time, and membership is anonymous (Oakes & Turner, 1980).

A second reason why people may have difficulties in accurately assessing their team's performance is that they establish high levels of identification with their teams during and after team tasks (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). Thus, when members highly identify with a team, they have a harder time psychologically distancing themselves from that team and its tasks than do those with lower identification (Libby, Shaeffer, & Eibach, 2009; Trope & Liberman, 2010). Psychological distance enables people to perceive the overall picture and take a global perspective, that is, "focusing on the forest," instead of focusing on details and situation-specific components of an incident, that is, "seeing the trees" (Liberman & Förster, 2009). Thus, individuals with greater psychological distance will remember and recall their

team discussions from a global and third person perspective as a more abstract and superordinate representation focusing on central features rather than on idiosyncratic information (Liberman & Trope, 2008; Trope & Liberman, 2010). This viewpoint, in turn, encourages greater adaptive self-reflection (Ayduk & Kross, 2010). Thus, psychological distance can enhance the accuracy of assessments of team performance because students are more likely to recall central features of the team discussion that are relevant to performance, such as if everyone contributed pieces of information to team discussions. Unimportant details, such as the order in which the students spoke or their initial preferences, are less salient. In contrast, a reduced psychological distance due to a stronger identification with the team will constrain accurate performance assessments.

These two explanations help us understand why students' performance assessments can be distorted. However, even if the correlation between objective and self-assessed performance cannot be expected to be high (Sitzmann et al., 2010), students will be able to develop a rough idea of how they have performed in the team task. Members will be able to remember some aspects of the task and the team interaction that indicate what went well and what did not. Thus, there likely is a positive relationship between the students' performance assessment and the team's objective performance. Consistent with other research showing a positive—albeit a small- or medium-sized effect (Dithurbide, Sullivan, & Chow, 2009; Jehn, Northcraft, & Neale, 1999)—relationship between objective and self-assessed team performance, we offer the following hypothesis as a baseline from which our other hypotheses are built:

Hypothesis 1: There will be a positive relationship between students' assessment of their team performance and their team's objective performance.

Although students' assessments of team performance can be biased due to the tendency to self-enhance or a strong identification with the team, the extent of this obstruction likely depends on the level of perceived relationship conflict within the team. *Relationship conflict* is defined as "interpersonal incompatibilities among group members, which typically includes tension, animosity, and annoyance among members within a group" (Jehn, 1995: 258). Thus, relationship conflict is related to interpersonal issues and is distinct from task conflict (i.e., conflict about the *content* of the team task and different task-related opinions). In the following, we will argue how perceived relationship conflict can mitigate the obstructions to students' accurate assessments of team performance.

Relationship conflict likely reduces students' self-enhancement tendencies. According to the depressive realism view (Alloy & Abramson, 1979), negative affect leads to less bias and more realistic information processing. As such, team members who perceive higher relationship conflict are likely to experience higher levels of negative affect (Jehn, 1995; von Glinow, Shapiro, & Brett, 2004), thereby increasing the accuracy of their team performance assessments. Furthermore, relationship conflict can lead students to develop a more negative view of the team (De Dreu & van Knippenberg, 2005), thus reducing in-group favoritism (Hogg & Terry, 2000). Because of their reduced tendency to self-enhance—based on higher levels of perceived relationship conflict—these students are likely to more accurately assess their team's performance than those who perceive lower levels of relationship conflict.

Furthermore, perceptions of higher relationship conflict will likely decrease an individual's identification with the team, thus will increase the psychological distance between the individual and the team (De Dreu & Beersma, 2005; De Dreu & van Vianen, 2001). Students who perceive relationship conflict to be higher will likely distance themselves from their team, as people have a tendency to distance themselves from others with characteristics that they judge to be undesirable (Schimel, Pyszczynski, Greenberg, O'Mahen, & Arndt, 2000). This psychological distance focuses the team members' attention on central features of the task such that they will have a more overall view of team performance (Liberman & Trope, 2008; Trope & Liberman, 2010). That is, when asked to assess their team's performance, students who have gained greater psychological distance as a result of perceived relationship conflict are in a more adequate "state of mind" to judge team discussions and actions holistically and to retrospectively acknowledge what went well and what did not. In contrast, when students perceive relationship conflict to be lower, they strongly identify with the team and have difficulties distancing themselves sufficiently to gain a global perspective about the team and its performance on the task; this will likely result in lower accuracy in their assessments of team performance. Thus,

Hypothesis 2: Relationship conflict, as perceived by individual students, will moderate the positive relationship between students' assessment of their team performance and their team's objective performance such that

higher relationship conflict will lead to a stronger assessed-objective performance relationship compared to lower relationship conflict.

Teams' Assessment of Team Performance

A student team's collective assessment of its performance refers to a process in which "teams receive a single questionnaire, members discuss their personal responses to each item, and they then decide by consensus which response best captures the overall team's response" (Quigley, Tekleab, & Tesluk, 2007: 590). In our study, the team's response refers to its subjective belief on how well the team worked in the team task (Dithurbide et al., 2009; Jehn, Chadwick, & Thatcher, 1997; Wittenbaum & Bowman, 2004). A student team's collective assessment of its performance will likely be related to its objective performance but will not necessarily perfectly reflect it. Teams generally want to achieve positive in-group evaluations (Brewer, 1991). Thus, the team as a whole will be inclined to increase its image by assessing its performance in a team-serving way (Aguinis, 2013). When students collectively discuss the assessment of team performance, they will likely place greater emphasis on those criteria in which the team excelled and less emphasis on those criteria in which it did not. Such a collective decision-making process helps contribute to a positive team climate, which most people strive for in team situations (Baumeister & Leary, 1995). This decision-making process is often connected to a team's feeling of invulnerability that it can achieve its goals and that it performs on a very high level (Lindsley, Brass, & Thomas, 1995). This will also lead to imprecise—more positive—assessments of team performance. Additionally, teams often create a climate where there is an expectation of conformity among members and minimal criticism directed at the team (Levine & Moreland, 1990; Postmes, Spears, & Cihangir, 2001). In particular, when cohesion in a team is high, criticism is usually suppressed (Postmes et al., 2001). This lack of criticism will also focus the team's assessment on the positive aspects of its performance but will, ultimately, reduce its assessment accuracy. Thus, these tendencies to "team enhance" and to limit criticism of the team likely distort the team's assessment of its performance.

After the team task, the common team identity will still be highly salient to the members, resulting in their high identification with the team (Lee, 2007; Mackie, 1986). This will reduce the team's psychological distance to team discussions and to the task at hand (cf. Libby et al., 2009). For example, people usually describe teams in which they are involved in a less abstract way than teams in which they are not involved because they feel less familiar with the latter and hence focus on general patterns describing them (Linville, Fischer, & Yoon, 1996). This identification with the team can obstruct an accurate assessment of team performance. However, despite obstructions, objective performance is likely to still influence the assessment of team performance. Thus, we offer the following as a second baseline hypothesis:

Hypothesis 3: There will be a positive relationship between a student team's assessment of team performance and the team's objective performance.

Teams that perceive higher relationship conflict will likely experience a more hostile communication during the task (Pelled, 1996). The team will be less inclined to collectively protect the team environment because relationship conflict signals that this climate is flawed. Thus, members feel less obligated to give positive comments about the team. Further, members are less motivated to "build up" the team climate because their need to belong to the team is likely reduced (Baumeister & Leary, 1995). This reduced need to belong decreases members' motivation for positive evaluations (cf. Brewer, 1991). Therefore, relationship conflict likely leads students to provide less biased, more accurate assessments of team performance.

Further, members feel less involved in teams that experience higher levels of relationship conflict (Hobman, Bordia, & Gallois, 2003). When individuals' involvement with their team is lower, they perceive the team in a more abstract and general way (Linville et al., 1996). Thus, relationship conflict will enhance the psychological distance of the team from its discussions and actions during its task. Because of this greater distance, the team will be better able to reflect about the task from a more neutral and global position (Ayduk & Kross, 2010; Liberman & Förster, 2009). Consequently, teams with higher levels of perceived relationship conflict will discuss performance at a more abstract level and in a "cool" and calculated fashion (Kross & Ayduk, 2008). This type of discussion will help student teams derive a more accurate image of their performance. In contrast, student teams with lower levels of perceived relationship conflict will be more involved with the team and their task, making it more difficult for them to establish psychological distance and, as a result, hampering them from seeing the big picture of their team's performance (Liberman & Trope, 2008; Trope & Liberman, 2010). Thus,

Hypothesis 4: Relationship conflict, as perceived by a student team, will moderate the positive relationship between a team's assessment of team performance and the team's objective performance such that higher relationship conflict will lead to a stronger assessed-objective performance relationship compared to lower relationship conflict.

METHODS

Participants, Research Setting, and Design

The sample consisted of 52 teams comprising 156 undergraduate business students enrolled at a German university. The students were recruited in business and economics courses to ensure that they could make an informed decision about an entrepreneurial opportunity (see description below). Students were compensated with €20 each (about US$25) for their participation. To minimize the effects of previous interactions, we randomly assigned students to teams of three members and invited them to a session. We asked each student how well they knew the other two students on 7-point Likert scales with the anchors "*I do not know him/her at all*" to "*I know him/her very well*." The participants indicated that they did not know the other members well ($M = 1.79$, $SD = 1.83$). On average, the participants were 24.31 years old ($SD = 2.54$) and 73 (46.8%) were male.

Our study consisted of implementing a decision-making task based on the hidden-profile technique (Stasser & Titus, 1985). Hidden profiles are situations where the best solution of a decision task is not initially evident to the team members from their personal information. Alone, a team member's information set points toward a suboptimal solution, but when all the unique information across individuals is pooled, then a best solution becomes obvious. Hidden-profile tasks have been used to investigate team decision making in social psychology (Schulz-Hardt et al., 2006; Stasser & Titus, 1985); organizational behavior (Alge, Wiethoff, & Klein, 2003; Okhuysen & Eisenhardt, 2002); and communication research (Cruz, Boster, & Rodriguez, 1997; Savadori, van Swol, & Sniezek, 2001). In our study, the teams were asked to choose the most attractive business opportunity from four alternatives. Before the team discussion, students received information sets about the decision alternatives. Some pieces of information were given to only one member and some pieces were given to all team members. These pieces of information were distributed among members in such a way that only when all the information was pooled by the members could the optimal alternative be identified.

Table 1 displays the distribution of the information pieces among the participants. Overall, there were 32 pieces of information—8 for each decision alternative business opportunity. For the optimal solution (Alternative A, Table 1), there are 6 positive and 2 negative pieces of information; whereas for the suboptimal alternatives (B–D, Table 1), there are 3 positive and 5 negative pieces of information. However, the individuals' information sets contained more positive than negative pieces of information for all suboptimal alternatives, but the

TABLE 1
Distribution of Information in the Hidden-Profile Decision-Making Task

	Alternative A	Alternative B	Alternative C	Alternative D
Common information				
Positive	0	3	3	3
Negative	2	0	0	0
Unique information				
Positive	6	0	0	0
Negative	0	5	5	5
Team member 1				
Positive	2	3	3	3
Negative	2	1	2	2
Team member 2				
Positive	2	3	3	3
Negative	2	2	1	2
Team member 3				
Positive	2	3	3	3
Negative	2	2	2	1

Note. Alternative A was the optimal solution.

same number of positive and negative pieces for the optimal solution. Thus, the optimal solution is hidden to the participants. Indeed, only 25 (16%) of our 156 participants chose the optimal alternative before the discussion based on their individual information set. However, in a pretest when all 32 pieces of information were available for an individual participant and participants were asked to rank the alternatives, 39 out of 45 students preferred the optimal solution, and it had a mean rank of 1.18. A Friedman test showed significant differences between the alternatives [$\chi^2(3) = 63.05$, $p < .000$, Kendall's $W = .43$] and subsequent paired Wilcoxon tests showed that the optimal alternative had a significantly better rank than all suboptimal alternatives (all z's > 5.35, $p < .001$) and the ranks of the suboptimal alternatives did not differ significantly from each other (all z's < 0.46, $p > .60$).

To avoid a potential confounding effect, we varied the order of ratings so that teams were randomly assigned to performance assessment from the individual's perspective first versus from the team's perspective first. We assigned the teams randomly to one of the two ordering conditions with the restriction that there were 26 teams in each condition.

Procedure

For each session, we invited three students to our lab. The experimenter welcomed them and informed them about the procedure of the study. The students were then asked to imagine being a part of an entrepreneurial team that has just invented a 3-dimensional printing technology (cf. Shane, 2000). They were told that they had already identified four potential business opportunities to exploit from their technology. As a team, they should now decide on one of these opportunities. Each of them would take a specific role on the entrepreneurial team—a marketing manager, a financial manager, and an operations manager. These managerial roles were randomly assigned, and each participant received an information set specific to the role. They were asked to study their information sets carefully to discuss the alternatives without needing to continuously check their sets, but they were allowed to keep their information sets during the subsequent team interaction. The participants had as much time as they needed to become familiar with the situation and their information sets.

After studying the materials, the teams were asked to discuss and decide which of the four alternatives they wanted to exploit as an entrepreneurial team of a new venture. We videotaped these discussions. Although the teams were told that they should take as much time as they needed, we also told them that teams usually finished within 30 min. We did not want to generate time pressure, but this time period was suggested to avoid "never-ending" discussions and to keep the teams focused on their task. When a team's discussion lasted longer than 30 min, we followed Schulz-Hardt and colleagues's (2006) procedure and reminded them of this time frame but did not specify further time limits. A discussion was considered to be finished when the team recorded its decision on a provided decision sheet. The average discussion time was approximately 22 min ($M = 21.63$ min, $SD = 8.14$ min).

After the discussion, the students filled out post-task questionnaires. For one half of the teams, we asked the students to individually fill out a questionnaire assessing team performance. After this individual assessment, we asked the team as a whole to assess the team's performance. The team was given a single copy of the same questionnaire and asked to come to a consensus with respect to each item (consensus method, cf. Quigley et al., 2007). They had as much time as they wanted to fill out the questionnaire. The other half of the teams were first asked to assess the team's performance from the team's perspective. Subsequently, the individual members were asked to assess the team's performance independently of the team's assessment.

After the performance assessments, we gave another questionnaire to the students individually. In this questionnaire, we assessed the team members' perceptions of task and relationship conflict. Subsequently, we captured the students' demographic details. Then, we debriefed the participants and revealed the nature of the hidden-profile task. Finally, we paid them for their participation in the study and they left the lab.

Measures and Variables

Dependent Variables: Individual and Team Assessments of Team Performance

To test our hypotheses, we specified two different models with different dependent variables—one at the individual and one at the team level. First, at the individual level, the dependent variable is each member's assessment of that team's performance. We used a 2-item scale based on Wittenbaum and Bowman (2004), including "*Our team performed well on the team task*" and "*Our team probably performed better on the team task than*

the average team in this study." A 7-point Likert scale with the anchors "*I do not agree at all*" and "*I completely agree*" was used to record self-assessed performance. We acknowledge that a 2-item scale is not ideal from a measurement theoretical perspective and that other measures have been applied in the assessment of student team performance (Kotey, 2007; Werner & Lester, 2001). However, this scale has already been applied in the context of a hidden-profile study (Wittenbaum & Bowman, 2004). Further, it captures general aspects of team performance which facilitates its application on the individual- and team-levels and does not relate to incidences that a team might not have experienced. In fact, the context of our study involving a general performance construct could possibly meet the criteria needed to use a 1-item scale (Wanous & Hudy, 2001). Second, at the team level, we used the teams' assessment of team performance as the dependent variable. We did this by using the same scale as for the individual ratings of team performance. Also, we used 7-point Likert scale to record the team-based assessments.

Objective Team Performance

Consistent with previous studies (Schulz-Hardt et al., 2006; Stasser & Titus, 1985) we defined *high team performance* as the team's choice for the best alternative in the hidden-profile task; whereas choosing a suboptimal alternative is defined as *low team performance*. This approach has the advantage of a clear difference between high and low objective team performance. Research on student teams often relies on grades assigned by instructors as an indicator of objective performance (Brown, 2003; Kotey, 2007; Werner & Lester, 2001). However, grading can be problematic, as it does not necessarily reflect objective performance (Roth, BeVier, Switzer, & Schippmann, 1996). In particular for team projects, the fairness of grades has been questioned because team members' effort and their contribution are difficult to assess (Baker, 2008; Brandyberry & Bakke, 2006; Paswan & Kamala, 2004; Willcoxson, 2006). Furthermore, these grades usually refer to the individual team member, even though the importance of assessing the performance of a team as a whole has been acknowledged (Brown, 2003; Goltz et al., 2008; Volkema, 2010). We directly derived objective team performance from the decision sheets that teams filled out at the end of the team task. We coded a decision as 1 when the team chose the best solution (i.e., Alternative A, Table 1), that is, the alternative with mainly positive pieces of information. All other decisions for suboptimal solutions (i.e., Alternatives B–D, Table 1), that is alternatives that had mainly negative pieces of information across all members' information sets, were coded as 0.

Perceived Relationship Conflict

We recorded the perceptions of relationship conflict during the team task using a scale developed by Jehn and colleagues (Jehn et al., 1997; Jehn, 1995). The scale consists of four items, including "How much interpersonal friction was there in your team?" Participants' answers were recorded on 7-point Likert scales with anchors "*not at all*" and "*very much*." Cronbach's alpha was .89 in our sample. For the individuals' perception of relationship conflict, we directly used the members' answers. For the teams' perception of relationship conflict, we averaged the values of each member within each team. Interrater reliability [*ICC* (1) = .40 and *ICC* (2) = .67] and interrater agreement (median $r_{wg(j)}$ = .95) indexes suggested that the members' perceptions were sufficiently similar to justify the aggregation of individual scores within teams (LeBreton & Senter, 2008).

Control Variables

As relationship conflict often follows or is accompanied by task conflict (Simons & Peterson, 2000), we controlled for perceived task conflict during the team task. By controlling for task conflict, we ensure that the personal friction within a team that helps members and teams to assess their performance more accurately is not the result of a different type of conflict related to the task. We used the scale developed by Jehn et al. (Jehn et al., 1997; Jehn, 1995), which consists of four items, including "How different were your views on the content of your project?" Participants' answers were recorded on 7-point Likert scales with anchors at "*not at all*" and "*very much*." The individuals' perceptions of task conflict were derived from the team members' answers. The members' values were then averaged for each team to obtain a team-level score of perceived task conflict. Acceptable values of interrater reliability [*ICC* (1) = .46 and *ICC* (2) = .72] and interrater agreement (median $r_{wg(j)}$ = .85) suggested that aggregation was justified (LeBreton & Senter, 2008).

Because students' performance likely depends on their task motivation (Latham & Pinder, 2005), we included a measure of this construct. We asked participants to assess their task motivation after the team task using a scale based on Colquitt and Chertkoff (2002). Our four items were "How motivated were you to exert effort in the team task?"

"How hard did you try to make a good selection?" "How much did you desire to make a good selection?" And, "How important was it for you to do well in the team task?" The answers were recorded on 7-point Likert scales with anchors "*not at all*" and "*very much*." To be able to control for task motivation also at the team level, we aggregated the individual values for each team.

As mentioned earlier, we randomized the order of the performance assessments. We controlled for the order of the performance assessment because individuals' assessments may be influenced by preceding teams' assessments and vice versa. This variable was dummy coded with 0 denoting individual assessments first and 1 denoting team assessments first.

As another control variable, we included the duration of the team task. Teams need time to exchange and process their information (Karau & Kelly, 1992) and a longer interaction could affect the team's performance and their assessment of their performance. This variable was derived from the videos of the interactions and was entered in hours in the analyses. To control for potential age and gender effects (cf. Riordan & Wayne, 2008) on the individual level, we asked students to indicate their year of birth and their gender at the end of our study. From the year of birth, we computed the students' age in years. Gender was entered as a dummy variable in the analysis, with 0 denoting males and 1 denoting females.

Translation Procedure

All scales were translated into German using a back-and-forth translation procedure recommended by Brislin (1970) to ensure maximum consistency between the translated and original scales. A native German speaker fluent in English translated the scales into German, and a native English speaker fluent in German translated it back to English. We compared the original and the back-translated versions and found no substantial differences.

Data Analysis

Testing our hypotheses necessitated two dependent variables, each relating to a different level of analysis. For Hypotheses 1–2 regarding the relationship of objective team performance and the individuals' assessments, we used a random coefficient modeling (RCM) approach as implemented by the software package Hierarchical Linear Modeling (HLM; Raudenbush & Bryk, 2002). This has the benefit that the nested structure of our data are taken into account (individuals are nested within teams) and that it enables us to analyze cross-level effects (i.e., the effect of a team-level independent variable on an individual-level dependent variable as in Hypothesis 1); and cross-level interactions (i.e., the moderating effect of an individual-level variable on the relationship between a team-level predictor variable and an individual-level dependent variable as in Hypothesis 2). All variables were grand-mean centered before the analyses. As an indicator for the explained variance in the dependent variable, we reported Pseudo R^2 based on the formula by Snijders and Bosker (1999). This statistic is based on the reduction of level 1 and level 2 residual variances (i.e., proportion of variance in the dependent variable left unexplained) after the predictors are included in the model.

For Hypotheses 3–4 regarding the relationship of objective team performance and the teams' performance assessment, we used ordinary least-squares (OLS) regression because the independent and dependent variables (assessment of team performance from the team's perspective) were all measured at the same (i.e., team) level of analysis. As independent variables, we included objective team performance and the team's perception of relationship conflict, and we computed and included an interaction term between these variables after centering them on their mean.

Although cautions have been issued about the use of control variables (Spector & Brannick, 2011), we included important ones for theory-based reasons and because they have been used as controls in related research (De Dreu, 2006; Jehn, 1995; Langfred, 2007). As a robustness check, we reran all analyses without control variables and found the same pattern of substantive results (i.e., direction, strength, and statistical significance of effects).

RESULTS

Table 2 includes the means, standard deviations, Cronbach's alphas, and correlations for all of the variables in this study. To test Hypotheses 1 and 2, we conducted RCM using individual assessments of team performance as the dependent variable. Results are displayed in Table 3. As a first step, we included the control variables (Model 1), which explained 8% of the variance of the individual students' performance assessments. Then, we included objective team performance and relationship conflict (Model 2) which accounted for 18% additional variance in individual performance assessments beyond that accounted for by the controls (total $R^2 = .26$). Further, Model 3 includes the

TABLE 2
Means, Standard Deviations, Cronbach's Alphas, and Correlations Between Study Variables

		M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12
1	Age[d]	24.31	2.54	(–)											
2	Gender[a, d]	0.53	0.50	.11	(–)										
3	Task motivation[d]	6.17	0.63	.01	−.04	(.68)									
4	Individual's perceived task conflict[d]	3.43	1.21	.09	.06	−.05	(.85)								
5	Individual's perceived relationship conflict[d]	1.66	0.97	−.01	−.03	−.04	.54**	(.89)							
6	Individual performance assessment[d]	5.18	0.96	.02	−.15	.20*	−.28**	−.27**	(.83)						
7	Order of assessments[b, e]	1.50	0.50	.08	.01	.04	−.06	−.05	.11	(–)					
8	Duration of discussion [h][e]	0.36	0.13	−.02	−.09	−.02	.36**	.23**	−.16	−.08	(–)				
9	Team task motivation[e]	6.17	0.40	−.01	−.15	.64**	−.17*	−.16	.19*	.05	−.04	(–)			
10	Team's perceived task conflict[e]	3.43	0.97	.13	.05	−.13	.80**	.48**	−.33**	−.07	.46**	−.21**	(–)		
11	Team's perceived relationship conflict[e]	1.66	0.75	.05	−.01	−.13	.49**	.77**	−.36**	−.06	.29**	−.20*	.62**	(–)	
12	Team performance assessment[e]	5.32	0.72	.03	−.20*	.20*	−.28**	−.34**	.65**	.09	−.15	.32**	−.35**	−.44**	(.72)
13	Objective team performance[c, e]	0.33	0.47	−.03	−.03	.00	−.16*	−.24**	.17**	−.29**	.05	.00	−.21**	−.31**	.29**

Note. $N = 156$ individuals, Cronbach's alpha (if applicable) is reported on the diagonal. M = mean; SD = standard deviation.
[a] 0 = "male," 1 = "female."
[b] 0 = "individuals' assessments first," 1 = "team's assessment first."
[c] 0 = "team chooses suboptimal solution," 1 = "team chooses optimal solution."
[d] These variables were measured at the individual level.
[e] These variables were measured at the team level of analysis and assigned down to individual team members for computing the correlations.
** $p < .01$. * $p < .05$.

TABLE 3
Random Coefficient Modeling Predicting Individual Assessments of Team Performance

Variables	Model 1 (control variables)		Model 2 (control variables and team-level objective performance)		Model 3 (control variables, team-level objective performance, and cross-level interaction effect)	
Level 1 first-order effects						
Intercept	5.23**	(0.09)	5.30**	(0.08)	5.36**	(0.07)
Age	−0.00	(0.0)	−0.02	(0.02)	−0.01	(0.02)
Gender[a]	−0.14	(0.13)	−0.15	(0.13)	−0.17	(0.12)
Task motivation	0.28**	(0.10)	0.34**	(0.10)	0.32**	(0.10)
Individual's perceived task conflict	−0.12	(0.07)	−0.08	(0.06)	−0.09	(0.06)
Individual's perceived relationship conflict			0.13	(0.08)	0.29**	(0.09)
Level 2 first-order effects						
Duration of discussion	−0.65	(0.69)	−1.37*	(0.61)	−1.28*	(0.57)
Order of performance assessments[b]	0.11	(0.20)	0.31	(0.17)	0.27	(0.15)
Objective team performance[c]			0.50**	(0.15)	0.65**	(0.14)
Cross-level interaction						
Individual's perceived relationship conflict × objective team performance					0.79**	(0.28)
Pseudo R^2	0.08		0.26		0.31	

Note. N = 156 individuals (level 1) in 52 teams (level 2).
Unstandardized estimates (based on grand-mean centering) are reported; robust standard errors are in parentheses.
Pseudo R^2 indicates the amount of total variance in the dependent variable explained by the predictors.
[a] 0 = "male," 1 = "female."
[b] 0 = "individuals' assessments first," 1 = "team's assessment first."
[c] 0 = "team chooses suboptimal solution," 1 = "team chooses optimal solution."
** $p < .01$. * $p < .05$.

cross-level effect between objective team performance and individuals' perceived relationship conflict. The cross-level interaction term accounted for 5% additional variance in individual performance assessments (total $R^2 = .31$).

Because we found a statistically significant cross-level effect, we interpret all lower order effects based on the full model including all predictors (i.e., Model 3, Aguinis, 2004). We found that objective team performance had a positive and significant effect on individual assessment of team performance ($\gamma = 0.65$, $p < .01$), which provides support for Hypothesis 1. The moderating effect of the individuals' perceived relationship conflict on the relationship between objective team performance and individuals' assessment of team performance was also positive and statistically significant ($\gamma = 0.79$, $p < .01$), providing support for Hypothesis 2. To better understand the nature of this cross-level interaction effect, we plotted this relationship in Figure 1. The *y*-axis represents individuals' performance assessments of team performance, and the *x*-axis is the objective performance. Following Aiken and West (1991) we plotted separate lines for higher (solid line, one standard deviation above the grand-mean of relationship conflict) and lower (dashed line, one standard deviation below the grand-mean of relationship conflict) levels of individuals' perception of relationship conflict. The slope is steeper (i.e., stronger) for higher levels of relationship conflict. In fact, $\gamma = 0.79$ means that a 1-point increase in relationship conflict is associated with a .79 in-

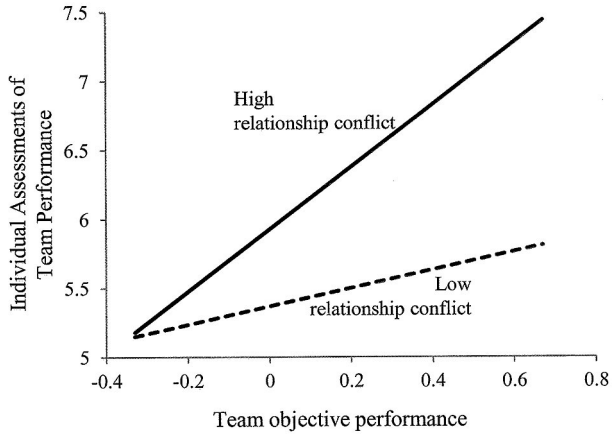

FIGURE 1
Moderating Effect of Relationship Conflict on the Relationship Between Individual Assessments of Team Performance and Objective Team Performance

crease in the slope of objective team performance on individual assessments of team performance.

Results regarding tests of Hypotheses 3 and 4 are included in Table 4. First, we included the control variables (Model 1); then, we added objective team performance and team's perceived relationship conflict (Model 2); finally, we included the product between objective performance and relationship conflict (Model 3). Model 1 led to a statistically significant $R^2_{adj} = .12$, $F(4,47) = 2.78$, $p < .05$. When objective team performance and the team's perceived relationship conflict are included (i.e., Model 2), explained variance increased to 21% [$F(6,45) = 3.21$; $p < .05$]. Next, we included the product term between objective team performance and the team's perceived relationship conflict (i.e., both team-level variables). The increase in explained variance was statistically significant, $\Delta R^2 = .07$, $p < .05$, $R^2_{adj} = .27$, $F(7,44) = 3.75$, $p < .001$. The coefficient for objective team performance is positive and statistically significant ($b = 0.72$; $p < .01$). Thus, Hypothesis 3 is supported. Further, the coefficient for the product term was positive and statistically significant ($b = 1.16$; $p < .05$). In other words, a 1-unit increase in relationship conflict is associated with a 1.16 increase in the effect of objective team performance on team assessments of team performance. To further probe this interaction, we plotted it in Figure 2. The *y*-axis represents the team's performance assessment, and the *x*-axis is the objective performance of the team. We plotted separate lines for higher (solid line, one standard deviation above the mean) and lower (dashed line, one standard deviation below the mean) levels of the teams' perceptions of relationship conflict. Figure 2 shows that the line for higher levels of relationship conflict is steeper, indicating that higher levels of relationship conflict are associated with a stronger relationship between objective performance and a team's assessment of its performance, which provides support for Hypothesis 4.

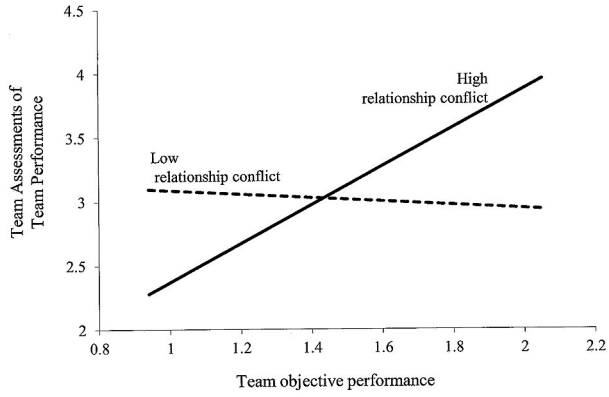

FIGURE 2
Moderating Effect of Relationship Conflict on the Relationship Between Team Assessment of Team Performance and Objective Team Performance

TABLE 4
Ordinary Least Square Regression Predicting Team Performance Assessments of Team Performance

	Model 1 (control variables)			Model 2 (control variables, objective team performance, and relationship conflict)			Model 3 (control variables, objective team performance, relationship conflict, and interaction effect)		
	b	*SE*	β	b	*SE*	β	b	*SE*	β
Intercept	3.24*	1.58		2.80	1.51		3.07*	1.45	
Order of performance assessments[a]	0.08	0.19	0.06	0.18	0.19	0.13	0.20	0.18	0.14
Duration of discussion	0.02	0.79	0.01	−0.16	0.76	−0.03	−0.26	0.73	−0.05
Task motivation	0.45	0.24	0.25	0.43	0.23	0.25	0.41	0.22	0.23
Team's perceived task conflict	−0.22	0.11	−0.30	−0.06	0.13	−0.08	−0.05	0.13	−0.07
Objective team performance[b]				0.36	0.21	0.24	0.72**	0.26	0.47
Team's perceived relationship conflict				−0.24	0.16	−0.25	0.07	0.20	0.08
Objective team performance × team's perceived relationship conflict							1.16*	0.51	0.43
Model fit	$R^2_{adj} = .12$; $F(4,47) = 2.78$*			$R^2_{adj} = .21$; $F(6,45) = 3.21$*; $\Delta R^2 = .10$**			$R^2_{adj} = .27$; $F(7,44) = 3.75$**; $\Delta R^2 = .07$**		

Note. $n = 52$.

[a] 0 = "individuals' assessments first," 1 = "team's assessment first."

[b] 0 = "team chooses suboptimal solution," 1 = "team chooses optimal solution." b = unstandardized regression coefficient, *SE*, β = standardized regression coefficient.

** $p < .01$. * $p < .05$.

DISCUSSION

Team projects are pervasive in all organizational settings, including higher education (Baldwin et al., 1997; Chen et al., 2004). We proposed a theory-based multilevel model that explains when students' self-assessments of team performance will be more or less accurate. We used a decision-making task and found support for our hypotheses. Results showed a positive relationship between a team's objective performance and self-assessments of team performance at the individual and team levels of analyses. Moreover, heterogeneity in these relationships can be explained by perceived relationship conflict on both levels—that is, higher levels of relationship conflict improve the accuracy of performance assessments both at the individual and the team levels. Interestingly, the duration of the discussion had a significant negative impact on the members' assessment of team performance. This indicates that the team members implicitly connected a slower team discussion to lower levels of team performance. Furthermore, the duration of the team discussion is positively correlated with task conflict. This reflects the conclusion by De Dreu (2008) that task conflict—which is often seen as positive for team outcomes—has costs for teams as it is connected to longer discussion which requires more time.

Implications for Research on Management in Education

Our study has implications for research on management education, in particular for research on self-assessments of performance. First, while many previous studies focused on individuals' self-assessments of performance (Schraw et al., 1993; Sitzmann et al., 2010), we included the social context—namely, other team members—in our analysis of the accuracy of self-assessments and found that the social context of these assessments indeed affects self-assessment accuracy. Because students are often nested in teams and classrooms, multilevel models are particularly suitable to analyze self-assessed performance because this approach allows researchers to specify the context in which the self-assessments were made. Our findings emphasize the importance of multilevel design, measurement, and analysis, and cover the individual and social contexts in understanding when self-assessments of performance are more or less accurate. Future research further exploring additional cross-level effects and considering the individual's (social or organizational) context could complement our understanding of the accuracy of self-assessed performance. Additionally, our results show that analyses focusing on different levels—the individual and team levels—can lead to parallel findings, which indicate that decision processes on both levels are isomorphic. Thus, future research could compare the processes of how self-assessments are made by comparing the team discussion about the performance assessment with an individual's assessment generated through a think-aloud technique. Testing these hypotheses regarding the role of context is only possible by way of multilevel design, measurement, and analysis tools.

Second, our results help educators to understand the role of relationship conflict in student team tasks. In educational contexts, relationship conflicts can be helpful for teams because they affect the accuracy of members' and teams' performance assessments. We argue that relationship conflict enables members and teams to be more accurate in their performance assessments because of the members' connected perspective. Future research should consider these perspectives when students are asked to assess their performance. This could help improve our understanding how performance assessments form and could also improve their accuracy.

Third, previous research has emphasized the role of self-enhancement theory (Barron & Sackett, 2008; Heidemeier & Moser, 2009; Klein, 2001) in explaining the process of how self-assessments develop. Complementing and expanding upon this literature, we draw on the construal level theory of psychological distance to suggest a process of how self-assessments form. While self-enhancement theory emphasizes the fact that students will assess their performance higher than it actually was, construal level theory builds on the notion that a greater distance between the assessor and the task is helpful for developing more accurate performance assessments. For the purpose of our study, construal level theory was particularly useful because it explains variance in the accuracy of performance assessments beyond self-enhancement tendencies. This helps us understand the heterogeneity in the relationship between self-assessed and objective performance, which can be attributed to the students' perspectives of the team task as being more or less distanced. These findings emphasize that scholars trying to understand student team processes in management education can profit from taking different theoretical perspectives when studying their phenomenon of interest.

Finally, Rynes and Brown (2011) investigated the legitimacy of the field of management education.

They argued that the field needs to be further advanced not only by thoroughly drawing on theory, but also by improving the methodologies used by management education scholars. Similar to Aguinis, De Bruin, Cunningham, Hall, Culpepper, and Gottfredson (2010), our study used a research method that is well established in other fields (e.g., psychology, statistics), but new to the management education literature. Specifically, we implemented a multilevel research design including a hidden-profile task to explore individual-level interactions in the team context and team-level objective outcomes. We have not created a new method but borrowed one to investigate a research question important to management education scholars. In doing so, we hope to have added a tool to the repertoire of methods for management education scholars. There are certainly limitations associated with using a hidden-profile task within a multilevel design in management education settings (as will be detailed below), but as we open ourselves to different methods, we also begin to ask different research questions. We hope others will use multilevel designs and hidden-profile studies to further our understanding of the processes, influences, and outcomes of student teams. Because research on management education often takes place in a context where a nested data structure is given, multilevel research could be particularly important for this field. Students are often grouped in teams, classrooms, or universities. Multilevel research enables researchers to specify the effects of different context factors and to test theories more adequately. So far research in management education has often neglected this data structure. For example, in the meta-analysis by Sitzmann and colleagues (2010), only a small minority of the published primary-level studies included measures at different levels. This will likely affect the studies' results because specific characteristics of the classroom, the teacher, or the learning environment are ignored. However, these characteristics can facilitate or impede the accuracy of self-assessments, as we have shown here. Thus, multilevel models can help to better understand phenomena in the field of management education and, at the same time, more complex methodologies will contribute to a greater legitimacy of the field.

Implications for Research on Assessments of Team Learning and Performance

Team members and teams need to develop an accurate self-assessment of their team's performance because this can provide them important corrective feedback (Aguinis, Gottfredson, & Joo, 2012; Schraw et al., 1993). However, previous research has shown that self-assessments are rather inaccurate in comparison with actual performance (Sitzmann et al., 2010). Our empirical findings illustrate that, at least for the specific situation of our empirical setting, the accuracy of self-assessments can be improved. Note that our model also acknowledges that the nature of self-assessed team performance can differ. That is, while previous studies usually rely on the individual team members' perspectives for self-assessments of team performance (Cohen & Bailey, 1997; Tziner & Eden, 1985), our multilevel hypotheses incorporate the members' and the teams' assessments of performance conjointly. By using a multilevel design, multilevel measurement, and multilevel analyses, we could show that the accuracy of assessments at both levels benefited from relationship conflict.

Further, our study extends research on team performance assessment by taking into account team interaction processes. Typically, work on the measurement of team performance has focused on explaining performance in terms of how satisfied the team is with the achieved outcome while neglecting that these assessments are influenced by interaction processes between team members (Jung & Sosik, 2002; Tziner & Eden, 1985; van Emmerik, 2008). Our results demonstrate that the assessment of team performance from the perspective of team members (either individually or collectively) is influenced by the nature of the process that led to this outcome, specifically the relationship conflict experienced. The implication of this finding is that studies investigating team performance should take into account factors of the team interaction process, and we encourage researchers to identify more characteristics of that process to help explain team members' performance assessments. For example, the communication quality of a team could be an interesting factor for future investigation. On the one hand, communication has been postulated to be crucial for student team performance (Hansen, 2006; Werner & Lester, 2001) and team members learn about each others' skills and expertise through high-quality communication (Lester, Meglino, & Korsgaard, 2002), which could improve the accuracy of their performance assessments. On the other hand, as high-quality communication is closely connected to cohesion (Barrick, Bradley, Kristof-Brown, & Colbert, 2007), it is also possible that it reduces the members' and team's psychological distance to the team task in a similar way as low levels of relationship conflict do. Thus, the inclusion of team interaction processes can help increase our understanding of how the members'

and the team's assessments of team performance form.

Implications for Research on Relationship Conflict

While much research on relationship conflict suggests that it has a negative impact on team performance (De Dreu & Weingart, 2003; Jehn, 1995; Langfred, 2007; Mohammed & Angell, 2004), our study challenges this finding by presenting a type of task for which this statement does not apply. Specifically, we show that relationship conflict improves the members' and team's ability to accurately assess team performance. The correlates and consequences of relationship conflict that usually impair team performance, such as a negative view of the team (De Dreu & van Knippenberg, 2005); hostile communication (Pelled, 1996); and lower levels of commitment to and identification with the team (De Dreu & Beersma, 2005; Hobman et al., 2003); seem to improve the accuracy of self-assessed team performance. This new facet of relationship conflict could be further tested in different assessment tasks. For example, perhaps relationship conflict will more generally enhance members' and teams' assessment accuracy—not only of team performance, but also of team processes and other nonperformance-related outcomes of these processes. We hope that our study points scholars into a new direction by stimulating research on the upside rather than on the downside of relationship conflict in teams.

In addition, our results suggest that students who perceive higher levels of relationship conflict can take a more global and abstract perspective and assess team performance in a more calculated (and thus objectively accurate) fashion. However, relationship conflict can also trigger strong negative affect, such as anger, frustration, and resentment (Jehn, 1995; Yang & Mossholder, 2004). This negative affect is connected to higher levels of arousal (Rozin & Royzman, 2001), which could inhibit a "cool" and calculated approach to performance assessment. Future research can make an interesting contribution by investigating what "states of mind" are connected to relationship conflict in team tasks and how these states of mind are reflected in the (individuals' or teams') self-assessments of performance.

Future research could further investigate how relationship conflict affects performance assessments. Identity could be an important moderator. Social identity theory (Hogg, 2001; Hogg & Terry, 2000) suggests that lower levels of identification could result in more inaccurate assessments because the team members reduce their involvement with the team (Hogg & Terry, 2000). Thus, they do not attach high importance to understanding the team processes (Hogg, 2001), which will reduce their ability to assess team performance. However, we theorized and found that relationship conflict—which is connected to lower levels of identification with the team (De Dreu & Beersma, 2005)—helps teams and their members to assess team performance more accurately. Thus, future research could focus on the different consequences of relationship conflict—such as negative affect or identification with the team—and how they help or impair accurate self-assessments.

Implications for Management Education Practice

Although our study has its focus on its methodological contribution, it also provides important implications for management education practice. An accurate self-assessment of performance is beneficial for both students and for management educators and—as the importance of teamwork increases in education (Boni et al., 2009; Chen et al., 2004)—the self-assessment of team performance is particularly relevant. For students, an accurate self-assessed team performance can represent corrective feedback to teams and help them to adjust and improve their performance (Schraw et al., 1993). Moreover, such feedback can be used by instructors to improve future student performance (Aguinis et al., 2012). In addition, a more accurate performance assessment will help adjust students' expectations about their grades and will, in consequence, reduce their disappointment with lower grades. Educators will also benefit from their students' more accurate performance assessments because it will help the students control the teamwork independent of the instructor. Additionally, educators will benefit from more appropriate student expectations because they will decrease the students' complaints about grading.[1]

Based on our results, educators can help teams develop such an accurate assessment by emphasizing the importance of the members' and the teams' distance from both the team and the team task when they assess team performance. For example, they could ask students to reflect about the team task from a third-person perspective, to focus on global elements of the team task, and to come to an abstract view of it (cf. Kross & Ayduk, 2008). This distance will enable students to more objectively

[1] We thank the participants of the *AMLE*'s writers workshop at the Academy of Management Meeting (2011) in San Antonio, TX, for pointing out this additional implication of the study.

and globally perceive their team's performance and to come to more accurate performance assessments. Thus, taking into account the context of a task and the assessor's state of mind can lead to more accurate self-assessments of performance, which is of benefit to individual students, student team, and educators.

In addition, our study shows that both the members' and the teams' assessments of team performance are enhanced by perceived relationship conflict. Thus, it appears that teams can rely on both types of assessments. Sometimes, it might be beneficial if the team as a whole discusses its performance. Then, the information and opinions of each member on team performance are accessible to the whole team, and members can learn from each other. On other occasions, it might be beneficial if the individuals reflect for themselves what went well and what did not. These reflections will not be distorted by the need to present oneself in a positive light in front of other members (Van Swol, 2009).

Limitations and Future Research

We readily acknowledge several limitations in our study that also provide the impetus for additional opportunities for future research. First, although we recorded self-assessed performance at the individual and team levels, we did not measure relationship conflict at both levels. Relationship conflict was measured at the individual level only and was then aggregated to a team-level construct. Interrater reliability and agreement provided support for aggregating the individual conflict scores to the team level, and this procedure is consistent with previous research on team conflict (De Dreu, 2006; Jehn, 1995; Mohammed & Angell, 2004). However, future research could additionally measure relationship conflict at the team level to differentiate between the individuals' and the teams' perceptions.

Second, objective team performance in our team task was dichotomous in nature. On the one hand, this is advantageous because there is a clear difference between high and low objective performance. On the other hand, our performance operationalization does not take into account that some teams that performed poorly came closer to the best solution than others. Thus, future research could rely on team tasks where there is more fine-grained measure of the teams' objective performance. Perhaps it is more difficult—or even easier—for members and teams to assess team performance in tasks without a clear "right or wrong" answer.

Third, to eliminate confounding influences of different team sizes, we restricted our teams to three members. This size is consistent with the literature on student teams, which typically describes team sizes between three to five members (Govekar & Rishi, 2007; Volkema, 2010; Werner & Lester, 2001). But we acknowledge that processes in teams can differ depending on team size (Bacon, 2005; Cohen & Bailey, 1997). Thus, future research could expand our decision-making task and include team size as an additional factor in the research design. As small teams often outperform larger teams in information exchange tasks (Cruz et al., 1997), smaller teams may be also better at assessing team performance.

Fourth, although we explicitly assessed the effects of age and gender (as control variables), demographic diversity can impact team processes and performance (Riordan & Wayne, 2008). In our study, participants were purposefully homogeneous regarding ethnicity as well as age. Because the main goal of our work was to study same as well as cross-level interaction effects, our research design did not include substantial heterogeneity regarding these demographic diversity variables that may interact with our hypothesized interaction effects (Riordan & Wayne, 2008). However, ethnic and age diversity are important issues to consider in future research. For example, whether our findings generalize to other samples should be tested, especially whether they generalize to students from other cultural or national backgrounds. This seems to be particularly interesting because Germans—coming from a rather individualistic culture (House, Hanges, Javidan, Dorfman, & Gupta, 2004)—are known to prefer a direct conflict style and do not strongly avoid impairments of social relationships (Bierbrauer & Klinger, 2005).

Finally, the team task in our study required that students interact in the laboratory setting for a short time only. Usually student teams will interact longer and will work on more complex tasks, which could impact the members' and the team's ability to assess team performance. In our study we used student teams that worked on a task for a short period; whereas most student teams work together for weeks (or even months). Further, we did not assess the members' teamwork skills and did not provide activities to enhance team development (Hansen, 2006). This represents a limitation of the current study because, for example, teams might develop group potency (Lester et al., 2002) or collective efficacy (Brown, 2003) with an increasing number of interactions, which has been found to impact team effectiveness, and which might also impact the accuracy of self-assessed team perfor-

mance. Indeed, issues of generalizability are a challenge for all laboratory studies (Shadish, Cook, & Campbell, 2002). However, we do note that our multilevel model was developed for student teams regardless of time together as a team although it was tested in a more restrictive context. Given that Mennecke and Valacich (1998) found that established teams are even worse than ad hoc teams in discussing unique information, our work here could be considered a more conservative test of our hypotheses. We hope that future research replicates the current study in teams formed for longer periods and theorizes on, and tests, the moderating role that "time as a team" has on the relationship between relationship conflict and the accuracy of self-assessed team performance. Additionally, we relied on undergraduate students as members of our student teams. In contrast to graduate students, undergraduates often have limited work experience and, thus, lack teamwork or project management skills (Volkema, 2010). This could affect their interaction processes and their performance assessments. Future research can investigate the potentially moderating role of teamwork experience or student level (undergraduate or graduate).

In closing, our results based on using multilevel design, measurement, and analysis and a hidden-profile task help scholars better understand self-assessments of performance in a team context and the particular role of relationship conflict for the accuracy of self-assessments. We emphasize that these theoretical contributions could be made because of the use of a new methodology in the field of management education and by explicitly taking into account the multilevel nature of student team interactions. This expansion of the methodological repertoire in the research on management education could also help future research to make important contributions.

REFERENCES

Aguinis, H. 2004. ***Regression analysis for categorical moderators.*** New York: Guilford.

Aguinis, H. 2013. ***Performance management*** (3rd ed.). Upper Saddle River, NJ: Prentice Hall/Pearson.

Aguinis, H., De Bruin, G. P., Cunningham, D., Hall, N. L., Culpepper, S. A., & Gottfredson, R. K. 2010. What does not kill you (sometimes) makes you stronger: Productivity fluctuations of journal editors. ***Academy of Management Learning and Education,*** 9: 683–695.

Aguinis, H., Gottfredson, R. K., & Joo, H. 2012. Delivering effective performance feedback: The strengths-based approach. ***Business Horizons,*** 55: 105–111.

Aguinis, H., Joo, H., & Gottfredson, R. K. 2011. Why we hate performance management—And why we should love it. ***Business Horizons,*** 54: 503–507.

Aiken, L. S., & West, S. G. 1991. ***Multiple regression: Testing and interpreting interactions.*** Thousand Oaks, CA: Sage Publications.

Alge, B. J., Wiethoff, C., & Klein, H. J. 2003. When does the medium matter? Knowledge-building experiences and opportunities in decision-making teams. ***Organizational Behavior and Human Decision Processes,*** 91: 26–37.

Alloy, L. B., & Abramson, L. Y. 1979. Judgment of contingency in depressed and nondepressed students: Sadder but wiser? ***Journal of Experimental Psychology: General,*** 108: 441–485.

Allport, G. W. 1937. ***Personality: A psychological interpretation.*** New York:. Holt.

Ayduk, O., & Kross, E. 2010. From a distance: Implications of spontaneous self-distancing for adaptive self-reflection. ***Journal of Personality and Social Psychology,*** 98: 809–829.

Bacon, D. R. 2005. The effect of group projects on content-related learning. ***Journal of Management Education,*** 29: 248–267.

Baker, D. F. 2008. Peer assessment in small groups: A comparison of methods. ***Journal of Management Education,*** 32: 183–209.

Baldwin, T. T., Bedell, M. D., & Johnson, J. L. 1997. The social fabric of a team-based M.B.A. program: Network effects on student satisfaction and performance. ***Academy of Management Journal,*** 40: 1369–1397.

Barrick, M. R., Bradley, B. H., Kristof-Brown, A. L., & Colbert, A. E. 2007. The moderating role of top management team interdependence: Implications for real teams and working groups. ***Academy of Management Journal,*** 50: 544–557.

Barron, L. G., & Sackett, P. R. 2008. Asian variability in performance rating—Modesty and leniency bias. ***Human Performance,*** 21: 277–290.

Baumeister, R. F., & Leary, M. R. 1995. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. ***Psychological Bulletin,*** 117: 497–529.

Bettencourt, B., & Dorr, N. 1998. Cooperative interaction and intergroup bias: Effects of numerical representation and cross-cut role assignment. ***Personality and Social Psychology Bulletin,*** 24: 1276–1293.

Bierbrauer, G., & Klinger, E. W. 2005. The influence of conflict context on conflict regulation preferences of immigrants. ***Journal of Cross-Cultural Psychology,*** 36: 340–354.

Boni, A. A., Weingart, L. R., & Evenson, S. 2009. Innovation in an academic setting: Designing and leading a business through market-focused, interdisciplinary teams. ***Academy of Management Learning and Education,*** 8: 407–417.

Bradley, G. W. 1978. Self-serving biases in the attribution process: A reexamination of the fact or fiction question. ***Journal of Personality and Social Psychology,*** 36: 56–71.

Brandyberry, A. A., & Bakke, S. A. 2006. Mitigating negative behaviors in student project teams: An information technology solution. ***Journal of Information Systems Education,*** 17: 195–209.

Brewer, M. B. 1991. The social self: On being the same and different at the same time. ***Personality and Social Psychology Bulletin,*** 17: 475–482.

Brislin, R. W. 1970. Back translation for cross-cultural research. ***Journal of Cross-Cultural Psychology,*** 1: 185–216.

Brown, T. 2003. The effect of verbal self-guidance training on collective efficacy and team performance. ***Personnel Psychology,*** 56: 935–964.

Chen, G., Donahue, L. M., & Klimoski, R. J. 2004. Training undergraduates to work in organizational teams. ***Academy of Management Learning and Education,*** 3: 27–40.

Cohen, S. G., & Bailey, D. E. 1997. What makes teams work: Group effectiveness research from the shop floor to the executive suite. ***Journal of Management,*** 23: 239–290.

Colquitt, J. A., & Chertkoff, J. M. 2002. Explaining injustice: The interactive effect of explanation and outcome on fairness perceptions and task motivation. ***Journal of Management,*** 28: 591–610.

Cruz, M. G., Boster, F. J., & Rodriguez, J. I. 1997. The impact of group size and proportion of shared information on the exchange and integration of information in groups. ***Communication Research,*** 24: 291–313.

De Dreu, C. K. W. 2006. When too little or too much hurts: Evidence for a curvilinear relationship between task conflict and innovation in teams. ***Journal of Management,*** 32: 83–107.

De Dreu, C. K. W. 2008. The virtue and vice of workplace conflict: Food for (pessimistic) thought. ***Journal of Organizational Behavior,*** 29: 5–18.

De Dreu, C. K. W., & Beersma, B. 2005. Conflict in organizations: Beyond effectiveness and performance. ***European Journal of Work and Organizational Psychology,*** 14: 105–117.

De Dreu, C. K. W., & van Knippenberg, D. 2005. The possessive self as a barrier to conflict resolution: Effects of mere ownership, process accountability, and self-concept clarity on competitive cognitions and behavior. ***Journal of Personality and Social Psychology,*** 89: 345–357.

De Dreu, C. K. W., & van Vianen, A. E. M. 2001. Managing relationship conflict and the effectiveness of organizational teams. ***Journal of Organizational Behavior,*** 22: 309–328.

De Dreu, C. K. W., & Weingart, L. R. 2003. Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. ***Journal of Applied Psychology,*** 88: 741–749.

Dithurbide, L., Sullivan, P., & Chow, G. 2009. Examining the influence of team-referent causal attributions and team performance on collective efficacy: A multilevel analysis. ***Small Group Research,*** 40: 491–507.

Freeman, L., & Greenacre, L. 2011. An examination of socially destructive behaviors in group work. ***Journal of Marketing Education,*** 33: 5–17.

George, J. M., & Bettenhausen, K. 1990. Understanding prosocial behavior, sales performance, and turnover: A group-level analysis in a service context. ***Journal of Applied Psychology,*** 75: 698–709.

Goltz, S. M., Hietapelto, A. B., Reinsch, R. W., & Tyrell, S. K. 2008. Teaching teamwork and problem solving concurrently. ***Journal of Management Education,*** 32: 541–562.

Goodman, P. S., & Leyden, D. P. 1991. Familiarity and group productivity. ***Journal of Applied Psychology,*** 76: 578–586.

Govekar, M. A., & Rishi, M. 2007. Service learning: Bringing real-world education into the b-school classroom. ***Journal of Education for Business,*** 83: 3–10.

Hansen, R. S. 2006. Benefits and problems with student teams: Suggestions for improving team projects. ***Journal of Education for Business,*** 82: 11–19.

Heidemeier, H., & Moser, K. 2009. Self-other agreement in job performance ratings: A meta-analytic test of a process model. ***Journal of Applied Psychology,*** 94: 353–370.

Hobman, E. V., Bordia, P., & Gallois, C. 2003. Consequences of feeling dissimilar from others in a work team. ***Journal of Business and Psychology,*** 17: 301–325.

Hogg, M. A. 2001. A social identity theory of leadership. ***Personality and Social Psychology Review,*** 5: 184–200.

Hogg, M. A., & Terry, D. J. 2000. Social identity and self-categorization processes in organizational contexts. ***Academy of Management Review,*** 25: 121–140.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. 2004. ***Culture, leadership, and organizations: The GLOBE study of 62 societies.*** Thousand Oaks: Sage Publications.

Jehn, K. A. 1995. A multimethod examination of the benefits and detriments of intragroup conflict. ***Administrative Science Quarterly,*** 40: 256–282.

Jehn, K. A., Chadwick, C., & Thatcher, S. M. B. 1997. To agree or not to agree: The effects of value congruence, individual demographic dissimilarity, and conflict on work group outcomes. ***International Journal of Conflict Management,*** 8: 287–305.

Jehn, K. A., Northcraft, G. B., & Neale, M. A. 1999. Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. ***Administrative Science Quarterly,*** 44: 741–763.

Jung, D. I., & Sosik, J. J. 2002. Transformational leadership in work groups: The role of empowerment, cohesiveness, and collective efficacy on perceived group performance. ***Small Group Research,*** 33: 313–336.

Karau, S. J., & Kelly, J. R. 1992. The effects of time scarcity and time abundance on group performance quality and interaction process. ***Journal of Experimental Social Psychology,*** 28: 542–571.

Klein, W. M. 2001. Post hoc construction of self-performance and other performance in self-serving social comparison. ***Personality and Social Psychology Bulletin,*** 27: 744–754.

Kotey, B. 2007. Teaching the attributes of venture teamwork in tertiary entrepreneurship programmes. ***Education and Training,*** 49: 634–655.

Kross, E., & Ayduk, O. 2008. Facilitating adaptive emotional analysis: Distinguishing distanced-analysis of depressive experiences from immersed-analysis and distraction. ***Personality and Social Psychology Bulletin,*** 34: 924–938.

Langfred, C. W. 2007. The downside of self-management: A longitudinal study of the effects of conflict on trust, autonomy, and task interdependence in self-managing teams. ***Academy of Management Journal,*** 50: 885–900.

Latham, G. P., & Pinder, C. C. 2005. Work motivation theory and research at the dawn of the twenty-first century. ***Annual Review of Psychology,*** 56: 485–516.

LeBreton, J. M., & Senter, J. L. 2008. Answers to 20 questions about interrater reliability and interrater agreement. ***Organizational Research Methods,*** 11: 815–852.

Lee, E.-J. 2007. Deindividuation effects on group polarization in computer-mediated communication: The role of group iden-

tification, public-self-awareness, and perceived argument quality. ***Journal of Communication,*** 57: 385–403.

Lester, S. W., Meglino, B. M., & Korsgaard, M. A. 2002. The antecedents and consequences of group potency: A longitudinal investigation of newly formed work groups. ***Academy of Management Journal,*** 45: 352–368.

Levine, J. M., & Moreland, R. L. 1990. Progress in small group research. ***Annual Review of Psychology,*** 41: 585–634.

Lewis, K., Belliveau, M., Herndon, B., & Keller, J. 2007. Group cognition, membership change, and performance: Investigating the benefits and detriments of collective knowledge. ***Organizational Behavior and Human Decision Processes,*** 103: 159–178.

Libby, L. K., Shaeffer, E. M., & Eibach, R. P. 2009. Seeing meaning in action: A bidirectional link between visual perspective and action identification level. ***Journal of Experimental Psychology: General,*** 138: 503–516.

Liberman, N., & Förster, J. 2009. Distancing from experienced self: How global-versus-local perception affects estimation of psychological distance. ***Journal of Personality and Social Psychology,*** 97: 203–216.

Liberman, N., & Trope, Y. 2008. The psychology of transcending the here and now. ***Science,*** 322: 1201–1205.

Lindsley, D. H., Brass, D. I., & Thomas, J. B. 1995. Efficacy-performance spirals: A multilevel perspective. ***Academy of Management Review,*** 20: 645–678.

Linville, P. W., Fischer, G. W., & Yoon, C. 1996. Perceived covariation among the features of ingroup and outgroup members: The outgroup covariation effect. ***Journal of Personality and Social Psychology,*** 70: 421–436.

Mackie, D. M. 1986. Social identification effects in group polarization. ***Journal of Personality and Social Psychology,*** 50: 720–728.

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. In press. Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. ***Journal of Applied Psychology***.

Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. 2008. Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. ***Journal of Management,*** 34: 410–476.

Mennecke, B. E., & Valacich, J. S. 1998. Information is what you make of it: The influence of group history and computer support on information sharing, decision quality, and member perceptions. ***Journal of Management Information Systems,*** 15: 173–197.

Mohammed, S., & Angell, L. C. 2004. Surface- and deep-level diversity in workgroups: Examining the moderating effects of team orientation and team process on relationship conflict. ***Journal of Organizational Behavior,*** 25: 1015–1039.

Oakes, P., & Turner, J. 1980. Social categorization and intergroup behaviour: Does minimal intergroup discrimination make social identity more positive? ***European Journal of Social Psychology,*** 10: 295–301.

Okhuysen, G. A., & Eisenhardt, K. M. 2002. Integrating knowledge in groups: How formal interventions enable flexibility. ***Organization Science,*** 13: 370–386.

Paswan, A. K., & Kamala, G. 2004. Dimensions of peer evaluation, overall satisfaction, and overall evaluation: An investigation in a group task environment. ***Journal of Education for Business,*** 79: 225–231.

Pelled, L. H. 1996. Demographic diversity, conflict, and work group outcomes: An intervening process theory. ***Organization Science,*** 7: 615–631.

Postmes, T., Spears, R., & Cihangir, S. 2001. Quality of decision making and group norms. ***Journal of Personality and Social Psychology,*** 80: 918–930.

Quigley, N. R., Tekleab, A. G., & Tesluk, P. E. 2007. Comparing consensus- and aggregation-based methods of measuring team-level variables: The role of relationship conflict and conflict management processes. ***Organizational Research Methods,*** 10: 589–608.

Rabbie, J. M., & Horwitz, M. 1969. Arousal of ingroup-outgroup bias by a chance win or loss. ***Journal of Personality and Social Psychology,*** 13: 269–277.

Raudenbush, S. W., & Bryk, A. S. 2002. ***Hierarchical linear models: Applications and data analysis methods*** (2nd ed.). Thousand Oaks: Sage Publications.

Riordan, C. M., & Wayne, J. H. 2008. A review and examination of demographic similarity measures used to assess relational demography within groups. ***Organizational Research Methods,*** 11: 562–592.

Roth, P. L., BeVier, C. A., Switzer, F. S., III, & Schippmann, J. S. 1996. Meta-analyzing the relationship between grades and job performance. ***Journal of Applied Psychology,*** 81: 548–556.

Rozin, P., & Royzman, E. B. 2001. Negativity bias, negativity dominance, and contagion. ***Personality and Social Psychology Review,*** 5: 296–320.

Rynes, S. L., & Brown, K. G. 2011. Where are we in the long march to legitimacy? Assessing scholarship in management learning and education. ***Academy of Management Learning and Education,*** 10: 561–582.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. 1992. Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), ***Teams: Their training and performance:*** 3–29. Norwood, NJ: Ablex.

Savadori, L., van Swol, L. M., & Sniezek, J. A. 2001. Information sampling and confidence within groups and judge advisor systems. ***Communication Research,*** 28: 737–771.

Schimel, J., Pyszczynski, T., Greenberg, J., O'Mahen, H., & Arndt, J. 2000. Running from the shadow: Psychological distancing from others to deny characteristics people fear in themselves. ***Journal of Personality and Social Psychology,*** 78: 446–462.

Schraw, G., Potenza, M. T., & Nebelsick-Gullet, L. 1993. Constraints on the calibration of performance. ***Contemporary Educational Psychology,*** 18: 455–463.

Schulz-Hardt, S., Brodbeck, F. C., Mojzisch, A., Kerschreiter, R., & Frey, D. 2006. Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. ***Journal of Personality and Social Psychology,*** 91: 1080–1093.

Shadish, W. R., Cook, T. D., & Campbell, D. T. 2002. ***Experimental and quasi-experimental designs for generalized causal inference.*** Boston: Houghton Mifflin.

Shane, S. 2000. Prior knowledge and the discovery of entrepreneurial opportunities. ***Organization Science,*** 11: 448–469.

Simons, T. L., & Peterson, R. S. 2000. Task conflict and relationship conflict in top management teams: The pivotal role of intragroup trust. ***Journal of Applied Psychology,*** 85: 102–111.

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. 2010. Self-assessment of knowledge: A cognitive learning or affective measure? ***Academy of Management Learning and Education,*** 9: 169–191.

Snijders, T. A. B., & Bosker, R. J. 1999. ***Multilevel analysis: An introduction to basic and advanced multilevel modeling.*** London: Sage.

Spector, P. E., & Brannick, M. T. 2011. Methodological urban legends: The misuse of statistical control variables. ***Organizational Research Methods,*** 14: 287–305.

Stasser, G., & Titus, W. 1985. Pooling of unshared information in group decision making: Biased information sampling during discussion. ***Journal of Personality and Social Psychology,*** 48: 1467–1478.

Stevens, M. J., & Campion, M. A. 1994. The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. ***Journal of Management,*** 20: 503–530.

Tonn, J. C., & Milledge, V. 2002. Team building in an MBA "gateway" course: Lessons learned. ***Journal of Management Education,*** 26: 415–428.

Trope, Y., & Liberman, N. 2010. Construal-level theory of psychological distance. ***Psychological Review,*** 117: 440–463.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. 1987. ***Rediscovering the social group. A self-categorization theory.*** New York: Basil Blackwell.

Tziner, A., & Eden, D. 1985. Effects of crew composition on crew performance: Does the whole equal the sum of its parts? ***Journal of Applied Psychology,*** 70: 85–93.

van Emmerik, I. J. H. 2008. It is not only mentoring: The combined influences of individual-level and team-level support on job performance. ***Career Development International,*** 13: 575.

Van Swol, L. M. 2009. Extreme members and group polarization. ***Social Influence,*** 4: 185–199.

Volkema, R. J. 2010. Designing effective projects: Decision options for maximizing learning and project success. ***Journal of Management Education,*** 34: 527–550.

von Glinow, M. A., Shapiro, D. L., & Brett, J. M. 2004. Can we talk, and should we? Managing emotional conflict in multicultural teams. ***Academy of Management Review,*** 29: 578–592.

Wanous, J. P., & Hudy, M. J. 2001. Single-item reliability: A replication and extension. ***Organizational Research Methods,*** 4: 361–375.

Werner, J. M., & Lester, S. W. 2001. Applying a team effectiveness framework to the performance of student case teams. ***Human Resource Development Quarterly,*** 12: 385–402.

Wilcoxson, L. E. 2006. "It's not fair!": Assessing the dynamics and resourcing of teamwork. ***Journal of Management Education,*** 30: 798–808.

Wilson, A. E., & Ross, M. 2001. From chump to champ: People's appraisals of their earlier and present selves. ***Journal of Personality and Social Psychology,*** 80: 572–584.

Wittenbaum, G. M., & Bowman, J. M. 2004. A social validation explanation for mutual enhancement. ***Journal of Experimental Social Psychology,*** 40: 169–184.

Yang, J., & Mossholder, K. W. 2004. Decoupling task and relationship conflict: The role of intragroup emotional processing. ***Journal of Organizational Behavior,*** 25: 589–605.

Nicola Breugst is a professor of entrepreneurial behavior at the Technische Universität München. She received her PhD in entrepreneurship at the University of Jena, Germany. Breugst's research focuses on entrepreneurial teams, and entrepreneurs' cognitive and affective processes.

Holger Patzelt is the Chair in Entrepreneurship and professor of entrepreneurship at the Technische Universität München. He earned a PhD in entrepreneurship at the Universität of Bamberg and a PhD in biosciences at the Universität of Heidelberg. His research focuses on entrepreneurial cognition and strategy in high-technology ventures.

Dean A. Shepherd is the Randall L. Tobias Chair in Entrepreneurial Leadership and professor of entrepreneurship at the Kelley School of Business, Indiana University. Shepherd earned his PhD at Bond University (Australia). His research focuses on opportunity, entrepreneurial decision-making, emotions, and failure.

Herman Aguinis is the Dean's Research Professor, a professor of organizational behavior and human resources, and the Director of the Institute for Global Organizational Effectiveness at the Kelley School of Business, Indiana University. His research and teaching interests are interdisciplinary and span several human resource management, organizational behavior, and research methods and analysis topics.